SPIN-2003/27
ITP-UU-03/44

3D Lorentzian Quantum Gravity from the asymmetric ABAB matrix model [1]

Jan Ambjørn[1,4], *Jerzy Jurkiewicz*[2], *Renate Loll*[3,4] *and Graziano Vernizzi*[3]

[1]The Niels Bohr Institute,
Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark
`email: ambjorn@nbi.dk`

[2]Inst. of Physics, Jagellonian University,
ul. Reymonta 4, 30-059 Krakow, Poland
`email: jurkiewicz@th.if.uj.edu.pl`

[3]Institute for Theoretical Physics, Utrecht University,
Leuvenlaan 4, NL-3584 CE Utrecht, The Netherlands
`email: r.loll@phys.uu.nl, g.vernizzi@phys.uu.nl`

[4]Perimeter Institute for Theoretical Physics,
35 King Street North, Waterloo, ON, Canada N2J 2W9
`email:jambjorn@perimeterinstitute.ca,`
`rloll@perimeterinstitute.ca`

Abstract

The asymmetric ABAB-matrix model describes the transfer matrix of three-dimensional Lorentzian quantum gravity. We study perturbatively the scaling of the ABAB-matrix model in the neighbourhood of its symmetric solution and deduce the associated renormalization of three-dimensional Lorentzian quantum gravity.

pacs: 04.60Gw, 04.20Gz, 04.60Kz, 04.60Nc

[1]Presented by J.J. at the Workshop on Random Geometry, May 15-17, 2003, Krakow, Poland

1 Introduction

Matrix models have been very useful in the study of the quantum geometry of two-dimensional quantum gravity. In [1] this program was extended to three-dimensional quantum gravity. It was shown how the so-called ABAB two-matrix model describes the transfer matrix of three-dimensional quantum gravity.[2] More precisely, a non-perturbative, background-independent definition of quantum gravity, which emphasizes the causal structure of space-time and which allows rotations between Lorentzian and Euclidean signature, was proposed in [2, 3], generalizing an explicitly solvable two-dimensional model with these features [4]. In the model, which has an UV lattice cut-off a which should be taken to zero in the continuum limit, one can define the concept of proper time. In the Euclidean sector the corresponding evolution operator is defined in terms of the transfer matrix $\hat{T}$ describing the transition between quantum states at (proper) time $n \cdot a$ and (proper) time $(n+1) \cdot a$. The transfer matrix is related to the quantum Hamiltonian of the system by

$$\hat{T} = \mathrm{e}^{-a\hat{H}}. \tag{1}$$

The ABAB model is defined by the two-matrix integral

$$\begin{aligned} Z(\alpha_1, \alpha_2, \beta) &= \mathrm{e}^{-M^2 F(\alpha_1, \alpha_2, \beta)} \\ &= \int \mathrm{d}A \mathrm{d}B \, \mathrm{e}^{-M \mathrm{tr}\,(\frac{1}{2}(A^2+B^2) - \frac{\alpha_1}{4}A^4 - \frac{\alpha_2}{4}B^4 - \frac{\beta}{2}ABAB)}. \end{aligned} \tag{2}$$

Under the assumption discussed in [1] the free energy $F(\alpha_1, \alpha_2, \beta)$ is related to the matrix elements of the transfer matrix $\hat{T}$ in a way reviewed in the next section. The matrix model (2) has a scaling limit for $\alpha_1 = \alpha_2$ which was analyzed in [7]. This allowed us in [1] to determine the corresponding phase diagram for the three-dimensional quantum gravity model and to map the bare coupling constants of the gravity model to the matrix model coupling constants $\alpha_1 = \alpha_2$ and β [9]. However, in order to study details of the scaling relevant to three-dimensional quantum gravity we have to study the matrix model for $\alpha_1 \neq \alpha_2$. In the scaling limit of interest for us both α_1 and α_2 will scale to a critical value α_c, but independently. Since we are interested only in the behaviour of the theory near the symmetric solution we need only the perturbative expansion around this solution rather than the complete solution in the asymmetric case[3].

The rest of this article is organized as follows. In sec. 2 we review shortly the non-perturbative definition of three-dimensional Lorentzian quantum gravity [2, 5] and its relation to the ABAB matrix model. In sec. 3 we review the machinery needed to

[2]Previous work on 3D quantum gravity in terms of Lorentzian triangulations can be found in [5, 6].

[3]While writing this article the asymmetric ABAB matrix model has been solved by Paul Zinn-Justin [10]. The behaviour close to the symmetric line $\alpha_1 = \alpha_2$ is the same as the one reported here and to extract it one has to expand the elliptic functions which appear in the solution, an effort comparable to the one used here.

solve the ABAB matrix model for a symmetric choice of coupling constants [7]. In Sec. 4 we discuss the solution of the general ABAB matrix model, and in Sec. 5 we expand around the symmetric critical point relevant for three-dimensional quantum gravity. In Sect. 6 we discuss how to extract information about the transfer matrix of 3D gravity, knowing the free energy of the asymmetric ABAB matrix model.

2 Quantum gravity and the ABAB matrix model

Simplicial Lorentzian quantum gravity in three dimensions is defined in the following way: the spatial hypersurfaces of constant proper-time are two-dimensional equilateral triangulations. Such triangulations define uniquely a two-dimensional geometry. It is known that this class of geometries describes correctly the quantum aspects of two-dimensional Euclidean gravity. It is also known that the description of two-dimensional Euclidean quantum gravity in terms of the class of (generalized) triangulations is quite robust. In [2] we used this universality in the following way: the two-dimensional geometry of the spatial hypersurfaces is represented by quadrangulations and it was shown that it is possible to connect any such pair of quadrangulations by a set of three-dimensional generalized "simplices". More precisely, let a be the lattice spacing separating two neighbouring spatial hypersurfaces at (proper)-times t and $t+a$. Then each square at t is connected to a vertex at $t+a$ and each square at $t+a$ is connected to a vertex at proper-time t, forming pyramids and inverted pyramids. A further needed three-dimensional building block is a tetrahedron connecting a spatial link at t to a spatial link at $t+a$. The proper-time propagator for (regularized) three-dimensional quantum gravity between two spatial hypersurfaces separated by a proper time $T = n \cdot a$ is obtained by inserting n–1 intermediate spatial hypersurfaces and summing over all possible geometries constructed as described above. The weight of each geometry is given by the Einstein action, here conveniently the Regge action for piecewise linear geometries. According to [3, 5], the contribution to the action from a single discrete time step is given by

$$S = -\kappa(N_t + N_{t+a} - N_{22}) + \lambda(N_t + N_{t+a} + \frac{1}{2}N_{22}), \tag{3}$$

where N_t, N_{t+a} and N_{22} denote the number of pyramids, upside-down pyramids and of tetrahedra contained in the slice $[t, t+a]$, and κ and λ are the dimensionless bare inverse gravitational and bare cosmological constant in three space-time dimensions. The naive continuum limit is obtained by scaling the lattice spacing $a \to 0$ while keeping $T = n \cdot a$ fixed. However, different scaling relations between T and a might in principle be possible[4].

[4] In two-dimensional *Euclidean* quantum gravity the proper-time T scales anomalously and one has to keep $n\sqrt{a}$ fixed. This is in contrast to the situation in two-dimensional *Lorentzian* quantum gravity as defined in [4] where the proper-time T scales canonically. The relation between the two models is well understood [11].

Let g_t and g_{t+a} be spatial two-geometries at t and $t{+}a$, i.e. two quadrangulations and let $\langle g_{t+a}|\hat{T}|g_t\rangle$ be the transition amplitude or proper time propagator from t to $t{+}a$. By definition, $\hat{T}$ is the transfer matrix in the sense of Euclidean lattice theory, and it satisfies the axioms of a transfer matrix [5]. In the case where the spatial topology is that of S^2 it was argued in [1] that the continuum limit could be obtained as the large N scaling limit of the matrix model (2). Let N_t and N_{t+a} denote the number of squares in the quadrangulations associated with g_t and g_{t+a}. The two-volumes of the corresponding geometries are thus $N_t a^2$ and $N_{t+a}a^2$, respectively, and the relation to $F(\alpha_1,\alpha_2,\beta)$ defined by (2) is

$$F(\alpha_1,\alpha_2,\beta) = \sum_{g_t,g_{t+a}} \mathrm{e}^{-z_t N_t(g_t) - z_{t+a}N_{t+a}(g_{t+a})} \langle g_{t+a}|\hat{T}|g_t\rangle, \tag{4}$$

where z_t and z_{t+a} are dimensionless boundary cosmological constants. The naïve relation between the matrix model coupling constants and the bare dimensionful gravitational and cosmological coupling constants $G^{(0)}$ and $\Lambda^{(0)}$ of three-dimensional gravity is

$$\alpha_1 = \mathrm{e}^{\kappa-\lambda-z_t}, \qquad \alpha_2 = \mathrm{e}^{\kappa-\lambda-z_{t+a}}, \qquad \beta = \mathrm{e}^{-(\frac{1}{2}\lambda+\kappa)}, \tag{5}$$

where

$$\kappa = \frac{a}{4\pi G^{(0)}}\left(-\pi + 3\cos^{-1}\frac{1}{3}\right), \quad \lambda = \frac{a^3\Lambda^{(0)}}{24\sqrt{2}\pi}, \quad z_t = a^2 Z_t^{(0)}. \tag{6}$$

In this paper we shall discuss the non-perturbative renormalization of the coupling constants. In principle we are interested in the limit $z_t = z_{t+a} = 0$, i.e. $\alpha_1 = \alpha_2$. However, in order to be able to extract the information about the scaling of the boundary cosmological constants we have to keep z_t and z_{t+a} different from zero at intermediate steps. Thus these *boundary cosmological constants* should be viewed as source terms for the boundary area operator.

3 The symmetric case: $\alpha_1 = \alpha_2 = \alpha$

Let us for later convenience shortly review the technique for solving the matrix model (2) used in [7] (based on earlier results [8]).

By a character expansion of the term $e^{\frac{1}{2}\beta M \mathrm{tr}\, ABAB}$ one can write

$$Z(\alpha_1,\alpha_2,\beta) \sim \sum_{\{h\}} \left(\frac{M\beta}{2}\right)^{\frac{1}{2}\sum h_i - \frac{M(M-1)}{4}} c_{\{h\}}\, R_{\{h\}}(\alpha_1) R_{\{h\}}(\alpha_2), \tag{7}$$

where the sum is over the representations of $GL(M)$, characterized by the shifted highest weights $h_i = m_i + M - i$, $(i = 1,\dots,M)$, where the m_i are the standard highest weights and where the large-M limit of the coefficient $c_{\{h\}}$ is

$$\log c_{\{h\}} = -\sum_i \frac{h_i}{2}\Big(\log\frac{h_i}{2} - 1\Big) - \frac{1}{2}\log\Delta(h), \qquad \Delta(h) = \prod_{i<j}(h_i - h_j). \tag{8}$$

Finally, if $\chi_{\{h\}}$ denotes the character associated with $\{h\}$,

$$R_{\{h\}}(\alpha) = \int \mathrm{d}A \ \chi_{\{h\}}(A) \ \exp M\Big(-\frac{1}{2}\mathrm{tr}\, A^2 + \frac{\alpha}{4}\mathrm{tr}\, A^4\Big). \tag{9}$$

It is now possible to perform a double saddle point expansion of (7) and (9). In order to describe the formalism let us introduce the notation

$$\Re f(z) \equiv \frac{f(z+i0)+f(z-i0)}{2}, \qquad \Im f(z) = \frac{f(z+i0)-f(z-i0)}{2}. \tag{10}$$

This notation is useful when $f(z)$ has cuts. The saddle point expansion assumes the existence of an eigenvalue density $\tilde{\rho}(\lambda)$, or equivalently a resolvent associated with the matrix integral (9):

$$\omega(\lambda) = \frac{1}{M}\sum_k \frac{1}{\lambda - \lambda_k}, \qquad -\pi i \tilde{\rho}(\lambda) = \Im\omega(\lambda), \tag{11}$$

and (after rescaling $h \to h/M$) a density of highest weights $\rho(h)$, or the corresponding "resolvent" $H(h)$:

$$H(h) = \frac{1}{M}\sum_k \frac{1}{h - h_k}, \qquad -\pi i \rho(h) = \Im H(h). \tag{12}$$

In [7] the double saddle point expansion is analyzed in the case $\alpha_1 = \alpha_2 = \alpha$. The density $\rho(h)$ was assumed to be different from zero in the interval $[0, h_2[$, and equal to 1 in the interval $[0, h_1]$, where $0 < h_1 < h_2$. Further, for a given eigenvalue distribution λ_k of the matrix A coming from the saddle point of (9) one can define a function $L(h)$, with same cut as H by

$$\Re L(h_j) = \frac{2}{M}\frac{\partial}{\partial h_j}\log \chi_{\{h\}}(A(\lambda_k)). \tag{13}$$

The analysis of [7] shows that $L(h) = H(h) + F(h)$ where $F(h)$ is analytic on the cut of $H(h)$ but has an additional cut $[h_3, \infty[$ where

$$2\Re L(h) = \log\frac{h}{\alpha} + H(h). \tag{14}$$

It can now be shown that the function $D(h) = 2L(h) - H(h) - 3\log h + \log(h - h_1)$ only has square root type cuts on $[h_1, h_2]$ and $[h_3, \infty[$ and on these cuts satisfies the following equations:

$$\Re D(h) = \log\frac{h - h_1}{\beta h^2}, \qquad h \in I_0 = [h_1, h_2] \tag{15}$$

$$\Re D(h) = \log\frac{h - h_1}{\alpha h^2}, \qquad h \in I_1 = [h_3, \infty[. \tag{16}$$

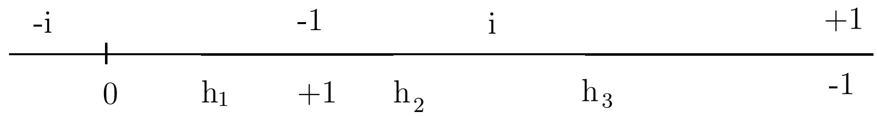

Figure 1: The cut structure of $r(h)$ in the complex h-plane.

Eqs. (15)–(16) constitute a standard Hilbert problem and the inversion formula is unique [12]. The function holomorphic in the plane with cuts I_0 and I_1 is given by

$$\begin{aligned} D(h) \;&=\; \log \frac{h-h_1}{\beta h^2} - \frac{\log \beta/\alpha}{i\pi}\, r(h) \int_{h_3}^{\infty} \mathrm{d}h' \, \frac{1}{(h-h')r(h')} \\ &\quad + r(h) \int_{-\infty}^{h_1} \mathrm{d}h' \frac{1}{(h-h')r(h')} - 2r(h) \int_{-\infty}^{0} \mathrm{d}h' \frac{1}{(h-h')r(h')} \end{aligned} \tag{17}$$

where

$$r(h) = \sqrt{(h-h_1)(h-h_2)(h-h_3)} \tag{18}$$

and where we have chosen the cut structure shown in Fig. 1. Following [12] the meaning of $r(h')$ on the cut is $r(h'+i0)$, i.e. the function on the "left side" of the cut. The integrals can be expressed in terms of standard elliptic functions. However, we do not need the explicit expressions here.

From $D(h)$ we can derive the expression for $\rho(h)$ which is

$$\rho(h) = -\frac{\Im H(h)}{i\pi} = -\frac{\Im D(h)}{i\pi}, \qquad h \in]h_1, h_2[. \tag{19}$$

We have (h always $h+i0$ if ambiguities) :

$$\begin{aligned} \rho(h) \;&=\; \frac{-r(h)}{\pi} \int_{-\infty}^{h_1} \frac{\mathrm{d}h'}{(h-h')ir(h')} - \frac{(-2r(h))}{\pi} \int_{-\infty}^{0} \frac{\mathrm{d}h'}{(h-h')ir(h')} + \\ &\quad + \frac{\log \beta/\alpha}{\pi^2} (-r(h)) \int_{h_3}^{\infty} \frac{\mathrm{d}h'}{(h-h')r(h')}. \end{aligned} \tag{20}$$

Note that the derivative of $D(h)$ and $\rho(h)$ with respect to h_i are elementary functions of h. For instance, differentiating with respect to h_3 we have

$$\frac{\partial D(h)}{\partial h_3} = \frac{ir(h)}{h-h_3}\,\frac{\partial W}{\partial h_3} = \frac{ir(h)F_3}{2(h-h_3)} = -i\pi \frac{\partial \rho(h)}{\partial h_3}, \tag{21}$$

where the last equality is valid for $h \in I_0 = [h_1, h_2]$ and where $W(h_1, h_2, h_3)$ and $F_3(h_1, h_2, h_3)$ are defined below (eqs. (25) and (28)). Thus we can write

$$\begin{aligned} D(h; h_3 + \delta h_3) \;&=\; D(h; h_3) + \delta h_3 \, \frac{ir(h)F_3(h_3)}{2(h-h_3)} \\ &\quad + (\delta h_3)^2 \frac{ir(h)}{4(h-h_3)} \left(F_3'(h_3) + \frac{F_3(h_3)}{2(h-h_3)} \right) + \cdots \end{aligned} \tag{22}$$

and similarly for $\rho(h)$. The function $F_3(h_1, h_2, h_3)$ is a sum of elliptic integrals.

3.1 Boundary conditions

The starting formula for $D(h)$ is (17). The general large-h behaviour of this function is

$$c_1 h^{1/2} - \log(-\alpha h) + c_2 h^{-1/2} + O(1/h). \tag{23}$$

However, according to the analysis in [7] $c_1 = 0$ and $c_2 = -(-\alpha)^{-1/2}$. This gives two boundary conditions for the constants h_1, h_2, h_3 which appear in $r(h)$ and thus in $D(h)$. The coefficients c_1 and c_2 can be identified by expanding the integrand in (17) in powers of $1/h$ and one obtains the boundary conditions:

$$c_1 = iW(h_1, h_2, h_3) = 0, \qquad c_2 = i\Omega = \frac{i}{\sqrt{\alpha}}, \tag{24}$$

where we have used the first of the equations (24) to simplify the second, and where

$$W(h_1, h_2, h_3) = \frac{\log \beta/\alpha}{\pi} \int_{h_3}^{\infty} \frac{\mathrm{d}h'}{r(h')} + \int_{-\infty}^{h_1} \frac{\mathrm{d}h'}{ir(h')} - 2\int_{-\infty}^{0} \frac{\mathrm{d}h'}{ir(h')}, \tag{25}$$

$$\Omega(h_1, h_2, h_3) = -\frac{\log \beta/\alpha}{\pi} \int_{h_1}^{h_2} \frac{h'\mathrm{d}h'}{r(h')} + \int_{h_2}^{h_3} \frac{h'\mathrm{d}h'}{ir(h')} + 2\int_{0}^{h_1} \frac{h'\mathrm{d}h'}{ir(h')}. \tag{26}$$

We will not need the explicit expressions for these integrals. The final boundary condition is

$$\int_{h_1}^{h_2} \mathrm{d}h\, \rho(h) = 1 - h_1, \tag{27}$$

where $\rho(h)$ is given by (20).

For given (α, β), we can in principle solve the three boundary conditions (24) and (27) for (h_1, h_2, h_3), thereby obtaining a solution of the matrix integral. Equivalently, the three conditions define locally a map from the β-α-plane to a two-dimensional hypersurface in the parameter space of the h_i. Critical regions of the free energy F of the model are associated with singularities of the inverse map of an independent subset of the h_i's, say, $(h_2, h_3) \mapsto (\alpha(h_2, h_3), \beta(h_2, h_3))$, which lead to singularities of $F(\alpha, \beta)$ upon substitution.

3.2 The critical line

The generic behaviour of $D(h)$ when $h \to h_3$ is clearly $(h - h_3)^{1/2}$, simply coming from the term $r(h)$ in the representation (17). However, this behaviour can change to $(h - h_3)^{3/2}$ along a curve $\alpha_c(\beta)$ in the (β, α) coupling constant plane. According to [7] this is the critical line of phase A of the ABAB matrix model and according to [1] this is where the continuum limit of 3d gravity should be found. Similarly the criticality in the B phase is derived from the behaviour of $D(h)$ when $h \to h_2$, where a generic behaviour $(h - h_2)^{1/2}$ changes to $(h - h_2)^{3/2}$.

We now study the change of (h_1, h_2, h_3) as α and β change infinitesimally. For simplicity we first present the result when α/β is constant.

Let us first identify the coefficient of $\sqrt{h-h_3}$ in $D(h)$. Using (21) in the expression for $D(h)$ the coefficient can be written as

$$ir_0(h_3)F_3(h_1,h_2,h_3), \qquad r_0(h) \equiv \sqrt{(h-h_1)(h-h_2)}, \tag{28}$$

One has

$$\frac{\partial W(h_1,h_2,h_3)}{\partial h_3} = \frac{1}{2}F_3(h_1,h_2,h_3). \tag{29}$$

We define F_1 and F_2 similarly to F_3 and have relations like (29). From the boundary conditions (24) it follows that the variation of h_1, h_2, h_3 as α, β change with the ratio α/β fixed satisfy

$$F_1\delta h_1 + F_2\delta h_2 + F_3\delta h_3 = 0. \tag{30}$$

$$h_1F_1\delta h_1 + h_2F_2\delta h_2 + h_3F_3\delta h_3 = -2\frac{\delta\alpha}{\alpha^{3/2}}. \tag{31}$$

The final boundary condition involves the density. Since $\rho(h_1)=1$ and $\rho(h_2)=0$ the variation of (27) just becomes

$$\int_{h_1}^{h_2} \mathrm{d}h \left(\frac{\partial\rho}{\partial h_1}\delta h_1 + \frac{\partial\rho}{\partial h_2}\delta h_2 + \frac{\partial\rho}{\partial h_3}\delta h_3\right) = 0, \tag{32}$$

which can be written, using (21), as

$$E_1F_1\delta h_1 + E_2F_2\delta h_2 + E_3F_3\delta h_3 = 0, \tag{33}$$

where

$$\int_{h_1}^{h_2} \frac{\mathrm{d}h\; r(h)}{h_i - h} = E_i, \qquad i = 1,2,3 \tag{34}$$

are elliptic integrals. It is straightforward to repeat the derivation in the case where the ratio α/β is not assumed constant, leading to the set of equations

$$\begin{aligned}
&R_1\left(\frac{\delta\alpha}{\alpha} - \frac{\delta\beta}{\beta}\right) + F_1\delta h_1 + F_2\delta h_2 + F_3\delta h_3 = 0,\\
&R_2\left(\frac{\delta\alpha}{\alpha} - \frac{\delta\beta}{\beta}\right) + \frac{\delta\alpha}{\alpha^{3/2}} + h_1F_1\delta h_1 + h_2F_2\delta h_2 + h_3F_3\delta h_3 = 0,\\
&R_3\left(\frac{\delta\alpha}{\alpha} - \frac{\delta\beta}{\beta}\right) + E_1F_1\delta h_1 + E_2F_2\delta h_2 + E_3F_3\delta h_3 = 0,
\end{aligned} \tag{35}$$

for the linear variations, where the functions $R_i(h_1,h_2,h_3)$ are given by

$$\begin{aligned}
&R_1 = -\frac{2}{\pi}\int_{h_3}^{\infty} \frac{dh'}{r(h')}, \quad R_2 = \frac{2}{\pi}\int_{h_1}^{h_2} \frac{h'dh'}{r(h')},\\
&R_3 = \frac{2}{\pi}\int_{h_1}^{h_2} dh\; r(h) \int_{h_3}^{\infty} \frac{dh'}{(h-h')r(h')}.
\end{aligned} \tag{36}$$

Without loss of generality, we can now choose h_2 and h_3 as independent variables on the 2d hypersurface. Using (35) to eliminate δh_1, one can derive the associated linear map between the remaining variables,

$$\begin{pmatrix} \delta\alpha/\alpha \\ \delta\beta/\beta \end{pmatrix} = \begin{pmatrix} X_1 F_2 & Y_1 F_3 \\ X_2 F_2 & Y_2 F_3 \end{pmatrix} \begin{pmatrix} \delta h_2 \\ \delta h_3 \end{pmatrix}, \tag{37}$$

where X_i and Y_i are easily calculable functions of the R's, E's and h's. This transformation is in general well defined, unless either $F_3 = 0$ or $F_2 = 0$, making the Jacobian vanish. We already know that these two equations define a critical line in coupling-constant space, the former giving rise to phase A, and the latter to phase B. They meet in a single critical point with $F_3 = F_2 = 0$.

In determining a continuum limit of the matrix model related to three-dimensional quantum gravity, we are interested in the relation between (β, α) and (h_2, h_3) along specific curves as they approach a point on the critical line. Note first that along generic curves and away from the critical line, by virtue of (37) all variations will be of the same order, namely

$$\delta\alpha \sim \delta\beta \sim \delta h_2 \sim \delta h_3. \tag{38}$$

This behaviour changes when the critical line is approached from an infinitesimal distance. As can be seen from (37), any such curve whose tangent does not coincide with that of the critical line in phase A at their mutual intersection point has vanishing derivatives $\partial\alpha/\partial h_3$ and $\partial\beta/\partial h_3$ there, by continuity. This means that although the variations $\delta\alpha$ and $\delta\beta$ are of the same order as δh_2 (and therefore also δh_1), their relation with δh_3 is of higher order, indeed,

$$\delta\alpha \sim \delta\beta \sim \delta h_2 \sim (\delta h_3)^2. \tag{39}$$

A completely analogous statement holds in phase B of the model with h_2 and h_3 interchanged, since the critical line is defined by $F_2 = 0$ there, instead of $F_3 = 0$. The qualitative difference between the two phases will only become apparent in the discussion of the asymmetric case below.

An alternative way of approaching a point on the critical line that turns out to be relevant for quantum gravity is by moving *along* the line itself. This case is analyzed most transparently by adding (in phase A) the constraint $F_3 = 0$ and expanding it along with (24) and (27). One verifies by a short calculation that the first-order variations in this case behave according to (38).

4 The asymmetric case $\alpha_1 \neq \alpha_2$.

As mentioned earlier, the construction of the transfer matrix requires that we perturb away from $\alpha_1 = \alpha_2$. Let us discuss the general structure of the matrix model with $\alpha_1 \neq \alpha_2$ (as explained above we will only need an infinitesimal perturbation away

from $\alpha_1 = \alpha_2$ in the continuum limit)[5]. The main difference in the analysis of the matrix model with $\alpha_1 \neq \alpha_2$ compared to the situation $\alpha_1 = \alpha_2 = \alpha$ is that the saddle point solution involves two eigenvalue densities $\tilde{\rho}_1(\lambda)$ and $\tilde{\rho}_2(\lambda)$ corresponding to the two one-matrix integrals (9) with $\alpha = \alpha_1$ and $\alpha = \alpha_2$. Similarly, we will have two functions $L_1(h)$ and $L_2(h)$ corresponding to (13) since the eigenvalue densities $\tilde{\rho}_1(\lambda)$ and $\tilde{\rho}_2(\lambda)$ appear via the matrix $A(\lambda)$ in (13). On other hand we have only one density $\rho(h)$ coming from the saddle point of (7). In order to solve the saddle point equations it is natural to follow the same strategy as in [7] and make an educated guess about the analytic structure of the functions involved and then show the self-consistency of the solution. Since we have two functions $L_i(h)$, associated with the same $\rho(h)$ but different $\tilde{\rho}(\lambda)$'s, and the cut of $L(h)$ from $[h_3, \infty[$ can be traced to the saddle point equation for $\tilde{\rho}(\lambda)$ (see [7] for a discussion), it is natural to assume that $L_1(h)$ and $L_2(h)$ have a cut from $[0, h_2]$ (with $\rho(h) = 1$ in $[0, h_1]$), and that they have separate cuts $[h_3^{(1)}, \infty[$ and $[h_3^{(2)}, \infty[$. In the case where $\beta = 0$ this structure is indeed realized.

We can now write down the generalization of (15)-(16)

$$
\begin{aligned}
\Re[D_1(h)] &= \log \frac{h - h_1}{h^2 \beta} + \Re[f(h)] \qquad \forall h \in I_0 \equiv [h_1, h_2] && (40)\\
\Re[D_1(h)] &= \log \frac{h - h_1}{h^2 \alpha_1} \qquad \forall h \in I_1 \equiv [h_3^{(1)}, \infty[&& (41)
\end{aligned}
$$

and

$$
\begin{aligned}
\Re[D_2(h)] &= \log \frac{h - h_1}{h^2 \beta} - \Re[f(h)] \qquad \forall h \in I_0 \equiv [h_1, h_2] && (42)\\
\Re[D_2(h)] &= \log \frac{h - h_1}{h^2 \alpha_2} \qquad \forall h \in I_2 \equiv [h_3^{(2)}, \infty[. && (43)
\end{aligned}
$$

In (40)-(43) the D's are related to the L's and the function H as below eq. (14), that is,

$$D_i(h) = 2L_i(h) - H(h) - 3\log h + \log(h - h_1), \qquad i = 1, 2, \tag{44}$$

where the subtractions of the log's are made to ensure that the functions D_i's have square root cuts. As in the case of a single α, we assume that $L_i(h) = F_i(h) + H(h)$ where $F_i(h)$ is analytic on the cut I_0 of $H(h)$.[6] The function

$$f(z) \equiv F_1(z) - F_2(z) = L_1(z) - L_2(z) \tag{45}$$

is at this point unknown, but we can write $\Re[f(z)] = f(z)$ on I_0. If we assume $f(z)$ is known on I_0, eqs. (40)-(41) and (42)-(43) are standard singular integral equations

[5] As already mentioned an explicit solution of the asymmetric ABAB model has been published while this manuscript was being completed [10].

[6] The F_i should not be confused with the functions of the same name defined in (29) above.

of the Hilbert type and can readily be solved and one can write

$$D_1(z) = D_1^{kz}(z) + r_1(z) \oint_{I_0} \frac{dt}{2\pi i} \frac{f(t)}{(z-t)r_1(t)} \tag{46}$$

$$D_2(z) = D_2^{kz}(z) - r_2(z) \oint_{I_0} \frac{dt}{2\pi i} \frac{f(t)}{(z-t)r_2(t)} \tag{47}$$

where $D_{1,2}^{kz}(z)$ are given by formula (17) with $(\alpha, h_3) = (\alpha_1, h_3^{(1)})$ and $(\alpha, h_3) = (\alpha_2, h_3^{(2)})$, respectively, and

$$r_i(h) = \sqrt{(h-h_1)(h-h_2)(h-h_3^{(i)})}. \tag{48}$$

Furthermore, we have

$$\Im[D_k(h)] = -i\pi\rho(h), \qquad k = 1, 2. \tag{49}$$

Therefore, the "imaginary" parts of eqs. (46) and (47) are

$$i\pi\rho(h) = i\pi\rho_1^{kz}(h) - r_1(h) \fint_{I_0} \frac{dt}{\pi i} \frac{f(t)}{(t-h)r_1(t)} \qquad \forall h \in I_0 \tag{50}$$

$$i\pi\rho(h) = i\pi\rho_2^{kz}(h) + r_2(h) \fint_{I_0} \frac{dt}{\pi i} \frac{f(t)}{(t-h)r_2(t)} \qquad \forall h \in I_0 \tag{51}$$

where $\fint$ is the principal value of the integral. Eqs. (50)–(51) determine $f(z)$ and $\rho(h)$ in terms of the densities ρ_1^{kz}, ρ_2^{kz} (eq. (57)), corresponding to $\alpha = \alpha_1$, $h_3 = h_3^{(1)}$ and $\alpha = \alpha_2$, $h_3 = h_3^{(2)}$. We can obtain an equation for $f(z)$ by subtracting (50) and (51), leading to

$$i\pi(\rho_1^{kz}(h) - \rho_2^{kz}(h)) = \fint_{I_0} \frac{dt\ f(t)}{\pi i(t-h)} \left(\frac{r_1(h)}{r_1(t)} + \frac{r_2(h)}{r_2(t)} \right). \tag{52}$$

4.1 Uniqueness of the solution

Let us discuss the solution of (52) which is a singular integral equation. In order to bring it into a standard form of singular integral equations, and for convenience of later applications, we introduce the notation

$$\Delta\rho^{kz}(h) \equiv \frac{i\pi}{2} \Big(\rho_1^{kz}(h) - \rho_2^{kz}(h) \Big), \tag{53}$$

and

$$(t-h)\, k(h,t) \equiv \frac{1}{2} \sqrt{\frac{t-h_3}{h-h_3}} \left(\sqrt{\frac{h-h_3^{(1)}}{t-h_3^{(1)}}} + \sqrt{\frac{h-h_3^{(2)}}{t-h_3^{(2)}}} - 2\sqrt{\frac{h-h_3}{t-h_3}} \right), \tag{54}$$

where from now on h_3 will always refer to the average

$$h_3 = \frac{1}{2}(h_3^{(1)} + h_3^{(2)}), \tag{55}$$

and the function $r(h)$ will refer to (18) with h_3 given by (55).

The function $k(h,t)$ is regular at $h{=}t$. Let us further introduce

$$\Delta\rho_r(h) \equiv \frac{\Delta\rho^{kz}(h)}{r(h)}, \qquad f_r(t) \equiv \frac{f(t)}{r(t)}. \tag{56}$$

We can now write eq. (52) as

$$-\!\!\!\!\!\!\int_{I_0} \frac{dt}{\pi i}\ \frac{f_r(t)}{t-h} + \int_{I_0} \frac{dt}{\pi i}\ k(h,t) f_r(t) = \Delta\rho_r(h), \tag{57}$$

where only the first integral is singular. The so-called *dominant part* of this singular integral equation is given by

$$-\!\!\!\!\!\!\int_{I_0} \frac{dt}{\pi i} \frac{\tilde{f}_r(t)}{t-h} = \Delta\rho_r(h). \tag{58}$$

This equation has precisely one zero mode, namely,

$$\tilde{f}_r(t) = \frac{1}{r_0(t)} -\!\!\!\!\!\!\int_{I_0} \frac{dh}{\pi i}\ \frac{r_0(h)\Delta\rho_r(h)}{h-t} + \frac{C}{r_0(t)}, \tag{59}$$

where $r_0(t) = \sqrt{(t-h_1)(t-h_2)}$. Expressed in terms of $\tilde{f}(t)$ we have

$$\tilde{f}(t) = \sqrt{t-h_3}\left(-\!\!\!\!\!\!\int_{I_0} \frac{dh}{\pi i}\ \frac{\Delta\rho^{kz}(h)}{(h-t)\sqrt{h-h_3}} + C \right). \tag{60}$$

By moving the k-term in Eq. (57) to the right-hand side, we can repeat the steps leading to (60), and moving the k-term back to the left-hand side we finally obtain

$$f(t) + \int_{I_0} \mathrm{d}s N(t,s)\, f(s) = \sqrt{t-h_3}\left(-\!\!\!\!\!\!\int_{I_0} \frac{dh}{\pi i} \frac{\Delta\rho^{kz}(h)}{(h-t)\sqrt{h-h_3}} + C \right), \tag{61}$$

where the kernel $N(t,s)$ is a *Fredholm kernel*,

$$N(t,s) = -\frac{\sqrt{t-h_3}}{r(s)} -\!\!\!\!\!\!\int_{I_0} \frac{\mathrm{d}h}{\pi^2}\ \frac{k(h,s) r_0(h)}{h-t}. \tag{62}$$

In general the solution to the Fredholm equation (61) will be unique [12]. *We thus have a one-parameter family of solutions $f_C(t)$.*

In order to determine the *four* parameters $h_1, h_2, h_3^{(1)}, h_3^{(2)}$ and the constant C we need *four* boundary conditions and one more condition, which in this case will

be the normalization condition for ρ (which appeared above in formulas (50) and (51)),

$$h_1 + \int_{h_1}^{h_2} \rho(t) dt - 1 = 0. \tag{63}$$

The boundary conditions are again obtained by the requirement that the large-h asymptotics of $D_j(h)$ contain no $h^{1/2}$ term while the coefficient of the $h^{-1/2}$ term equal $(-\alpha_j)^{-1/2}$. This results in four boundary conditions

$$\mathcal{W}_j := W_j - (-1)^j \int_{I_0} \frac{dt}{\pi} \frac{f(t)}{r_j(t)} = 0, \tag{64}$$

$$\mathcal{O}_j := \Omega_j - (-1)^j \int_{I_0} \frac{dt}{\pi} \frac{f(t)t}{r_j(t)} - \frac{1}{\sqrt{\alpha_j}} = 0, \tag{65}$$

where $j = 1,2$, $W_j = W(h_1, h_2, h_3^{(j)})$ and $\Omega_j = \Omega(h_1, h_2, h_3^{(j)})$.

The set of constraints (63) (64), (65) implies that in the asymmetric case the theory is defined on a three-dimensional hypersurface of the eight-dimensional parameter space $(\alpha_1, \alpha_2, \beta, C, h_1, h_2, h_3^{(1)}, h_3^{(2)})$. Singular points of the map between the parameters $(\alpha_1, \alpha_2, \beta)$ and some other parametrization of this hypersurface (say, in terms of $(h_2, h_3^{(1)}, h_3^{(2)})$, after eliminating C and h_1) will correspond to critical points of the theory.

In analogy with the symmetric case we expect criticality to be related to a change in the behaviour of the $D_j(h)$ when h approaches one of the $h_3^{(j)}$ (phase A) or h_2 (phase B). In the first case, this is signalled by the vanishing of one of the coefficients $\mathcal{F}_3^{(j)}$ in the relevant expansion

$$D_j(h) = \mathcal{F}_3^{(j)} \sqrt{h - h_3^{(j)}} + \text{higher order}, \tag{66}$$

with

$$\mathcal{F}_3^{(j)} := F_3^{(j)} - \int_{I_0} \frac{dt}{\pi} \frac{f(t)}{(h_3^{(j)} - t) r_j(t)} \tag{67}$$

where $F_3^{(j)} = F_3(h_3 = h_3^{(j)})$. The two critical hypersurfaces in phase A defined by $\mathcal{F}_3^{(j)}$, $j = 1, 2$, intersect along a critical line which coincides with the critical line of the symmetric model.

In phase B, there is a single smooth critical hypersurface defined by the simultaneous vanishing of the two functions $\mathcal{F}_2^{(j)}$ (defined similar to (67)) appearing in the expansions

$$D_j(h) = \mathcal{F}_2^{(j)} \sqrt{h - h_2} + \text{higher order}. \tag{68}$$

The simultaneous change in behaviour of $\mathcal{F}_2^{(1)}$ and $\mathcal{F}_2^{(2)}$ can be traced to the fact that the non-analyticity at h_2 is due entirely to $H(h)$ (or $\rho(h)$) which are common to $D_1(h)$ and $D_2(h)$. The critical hypersurface defined by $\mathcal{F}_2^{(j)} = 0$ intersects both

hypersurfaces $\mathcal{F}_3^{(1)}=0$ and $\mathcal{F}_3^{(2)}=0$ in critical lines, and there is a single (tricritical) point where all of the three hypersurfaces meet.

One can now proceed as in the symmetric case by expanding the five constraint equations to first (or any desired) order, and eliminating the dependent variations δh_1 and δC, say. The three remaining first-order equations can be solved for $\delta\alpha_1$, $\delta\alpha_2$ and $\delta\beta$ and put in matrix form,

$$\begin{pmatrix} \delta\alpha_1/\alpha_1 \\ \delta\alpha_2/\alpha_2 \\ \delta\beta/\beta \end{pmatrix} = \begin{pmatrix} \mathcal{X}_{11}\mathcal{F}_3^{(1)} + \mathcal{Z}_{11} & \mathcal{X}_{12}\mathcal{F}_3^{(2)} + \mathcal{Z}_{12} & \mathcal{X}_{13}\mathcal{F}_2^{(1)} + \mathcal{Z}_{13} \\ \mathcal{X}_{21}\mathcal{F}_3^{(1)} + \mathcal{Z}_{21} & \mathcal{X}_{22}\mathcal{F}_3^{(2)} + \mathcal{Z}_{22} & \mathcal{X}_{23}\mathcal{F}_2^{(1)} + \mathcal{Z}_{23} \\ \mathcal{X}_{31}\mathcal{F}_3^{(1)} + \mathcal{Z}_{31} & \mathcal{X}_{32}\mathcal{F}_3^{(2)} + \mathcal{Z}_{32} & \mathcal{X}_{33}\mathcal{F}_2^{(1)} + \mathcal{Z}_{33} \end{pmatrix} \begin{pmatrix} \delta h_3^{(1)} \\ \delta h_3^{(2)} \\ \delta h_2 \end{pmatrix}, \quad (69)$$

where the functions $\mathcal{X}_{ij}$ and $\mathcal{Z}_{ij}$ are regular and generically non-vanishing. Moreover, it can be shown that the three functions $\mathcal{Z}_{i1}$ vanish at points $h=h_3^{(1)}$ of the critical surface defined by $\mathcal{F}_3^{(1)}=0$, by virtue of the fact that the derivative $\partial f/\partial h_3^{(1)}$ at such points is proportional to the zero-mode of f. (Analogous statements hold for $\mathcal{Z}_{i2}$ and $\mathcal{Z}_{i3}$.) The situation is therefore very similar to what happened in the symmetric case, cf. (37). Namely, the rank of the Jacobian of the coordinate transformation $(\alpha_1,\alpha_2,\beta)\mapsto(h_3^{(1)},h_3^{(2)},h_2)$ drops by one when we go to one of the critical surfaces.

Expressed in terms of the linear variations, this means that away from any of the critical surfaces one has

$$\delta\alpha_1 \sim \delta\alpha_2 \sim \delta\beta \sim \delta h_2 \sim \delta h_3^{(1)} \sim \delta h_3^{(2)}, \quad (70)$$

and that this behaviour changes when a critical surface, say, $\mathcal{F}_3^{(1)}=0$ is approached, to

$$\delta\alpha_1 \sim \delta\alpha_2 \sim \delta\beta \sim \delta h_2 \sim (\delta h_3^{(1)})^2 \sim \delta h_3^{(2)}. \quad (71)$$

Similarly, it follows that if the critical (symmetric) line of phase A is approached along any path that is not tangential to either of the critical surfaces $\mathcal{F}_3^{(j)}$, $j=1,2$, the variations will behave like

$$\delta\alpha_1 \sim \delta\alpha_2 \sim \delta\beta \sim \delta h_2 \sim (\delta h_3^{(1)})^2 \sim (\delta h_3^{(2)})^2. \quad (72)$$

As before in the symmetric case, if one chooses a linear variation tangential to one of the critical hypersurfaces, say, $\mathcal{F}_3^{(1)}=0$, this imposes a linear relation between $\delta h_3^{(1)}$, $\delta h_3^{(2)}$ and δh_2, and at the same time forces $\delta h_3^{(1)}$ to be of the same order as the other two, leading again to the behaviour (70). The same is true if we consider a variation tangent to the critical line in phase A. This case is analyzed easily by adjoining the two additional conditions $\mathcal{F}_3^{(j)}=0$ to the other boundary conditions, and expanding them to first order.

5 Expanding around $\alpha_1 = \alpha_2$

In applying the matrix model to 3d quantum gravity, we do not need the complete explicit solution of the asymmetric ABAB model, but only infinitesimal variations

around symmetric points with $\alpha_1 = \alpha_2$ and $h_3^{(1)} = h_3^{(2)}$. In this context, it is convenient to work with the symmetrized and anti-symmetrized variables

$$\alpha = \sqrt{\alpha_1 \alpha_2}, \quad \tilde{\alpha} = \sqrt{\frac{\alpha_1}{\alpha_2}}, \quad h_3 = \frac{h_3^{(1)} + h_3^{(2)}}{2}, \quad \tilde{h}_3 = \frac{h_3^{(1)} - h_3^{(2)}}{2}. \tag{73}$$

Infinitesimal linear variations around any given point $(\alpha^{(0)}, \tilde{\alpha}^{(0)}, \beta^{(0)})$ or the corresponding $(h_3^{(0)}, \tilde{h}_3^{(0)}, h_2^{(0)})$ will be denoted by $\delta\alpha$, $\delta\tilde{\alpha}$, $\delta\beta$, δh_3, $\delta\tilde{h}_3$ and δh_2. In what follows, we will study further the relations between these variations which follow from the conditions (63), (64) and (65), since they determine the continuum physics of the model.

5.1 Finding $f(t)$ and $D_{1,2}(h)$

The discussion in this section will concentrate on phase A, which is the most relevant one for the quantum-gravitational application. We will comment briefly at the end on what happens in phase B.

In order to make maximal use of the relations we already derived for the symmetric case, it is convenient to perform a partial expansion of the relevant functions around points in the symmetric plane (characterized by $\tilde{\alpha} = 1$, or, equivalently, $\tilde{h}_3 = 0$). We will Taylor-expand around $\tilde{h}_3^{(0)} = 0$ and simultaneously allow for a finite shift to $\tilde{\alpha} = \mathrm{e}^{\varepsilon}$ away from the symmetric value $\tilde{\alpha}^{(0)} = 1$.

Recall now the integral equation (61) which determines $f(t)$. Under the variation introduced in the last paragraph, because of symmetry in $h_3^{(1)}$ and $h_3^{(2)}$, the function $k(t, h)$ has an expansion in even powers of $\delta\tilde{h}_3$,

$$k(t, h) = (\delta\tilde{h}_3)^2 k_1(t, h) + (\delta\tilde{h}_3)^4 k_2(t, h) + \cdots . \tag{74}$$

Thus knowing $\Delta\rho^{kz}(h)$ allows us to calculate $f(t)$ perturbatively in $\delta\tilde{h}_3$. We can also expand $\Delta\rho^{kz}(h)$ in $\delta\tilde{h}_3$ and ε. This makes the integration on the left-hand side of eq. (61) possible order by order in terms of elementary functions. The expansion of $\Delta\rho^{kz}(h)$ is based on the following two observations. Firstly, we have for any (not just infinitesimal) ε the exact relation

$$D^{kz}(h; \alpha\mathrm{e}^{\varepsilon}) = D^{kz}(h; \alpha) + \varepsilon\, \mathcal{G}(h), \tag{75}$$

$$\mathcal{G}(h) = \frac{1}{i\pi}\, r(h) \int_{h_3}^{\infty} \mathrm{d}h' \, \frac{1}{(h - h')r(h')}. \tag{76}$$

Next, $D^{kz}(h; h_3)$ and $\mathcal{G}(h)$ have expansions of the form (22) with respect to their h_3-dependence. Thus $\Delta\rho^{kz}(h)$ can be written as

$$\Delta\rho^{kz}(h; \delta\tilde{h}_3, \varepsilon) = (a_1\, \delta\tilde{h}_3 + b_1\, \varepsilon) + (a_2\, \delta\tilde{h}_3 + b_2\, \varepsilon)(\delta\tilde{h}_3)^2 + \cdots , \tag{77}$$

where only b_1 is not an elementary function of h. Explicitly, we have

$$\Delta\rho^{kz}(h) = \left(-\delta\tilde{h}_3 \frac{ir(h)}{2(h-h_3)} F_3(h_3) - \varepsilon\,\mathcal{G}(h,h_3)\right) + \cdots . \tag{78}$$

Here and in the following we use the shorthand notation $F_3(h_3)$ for the function $F_3(h_1, h_2, h_3, \alpha, \beta)$. The integral in (61) can now be performed and one obtains

$$f(t) = \left(C'\sqrt{t-h_3} + \delta\tilde{h}_3 \frac{r_0(h_3)}{\sqrt{t-h_3}} F_3(h_3) - \varepsilon\right) + \cdots , \tag{79}$$

where $C' = C + \delta\tilde{h}_3\, F_3(h_3) + \varepsilon \frac{1}{i\pi}\int_{h_3}^{\infty} \frac{dh'}{r(h')}$. Note that until this point we have not made use of any of the boundary conditions, and therefore could treat $\delta\tilde{h}_3$ and ε as independent.

Finally we can insert $f(t)$ in (46) and (47) to obtain $D_{1,2}(h)$. To order $(\delta\tilde{h}_3)^2$ we obtain

$$\begin{aligned} D_j(h; h_1, h_2, h_3^j; \alpha_j, \beta) &= D^{kz}(h; h_1, h_2, h_3; \alpha, \beta) + \\ &\mp\varepsilon \pm C'\left(\sqrt{h-h_3} + O(\delta\tilde{h}_3)\right) \pm \delta\tilde{h}_3 \frac{ir_0(h_3)F_3(h_3)}{2\sqrt{h-h_3}} \\ &+ \frac{(\delta\tilde{h}_3)^2}{4} \frac{ir_0(h)}{\sqrt{h-h_3}} \left(\frac{\partial F_3(h_3)}{\partial h_3} + \frac{F_3(h_3)}{2(h-h_3)} + \frac{F_3(h_3)}{2(h_3-h_1)} + \frac{F_3(h_3)}{2(h_3-h_2)})\right) \cdots \end{aligned} \tag{80}$$

Note first that $C'\!=\!0$ since the large-h asymptotics

$$D_j(h) \sim i\mathcal{W}_j h^{1/2} - \log\alpha_j h + i\mathcal{O}_j h^{-1/2} + O(1/h), \tag{81}$$

must be satisfied for *both* D_j's to zeroth (and, of course, to any) order in $\delta\tilde{h}_3$. Furthermore, note that the ε-dependence to this order is simply due to a shift in the argument from $\log\alpha$ to $\log\alpha_j\!=\!\log\alpha\,\pm\varepsilon$ in the terms $\log((h-h_1)/(h^2\alpha_j))$ in D_j.

In order to give a more detailed discussion of the continuum limit, we will also need the explicit form of the boundary conditions (63), (64) and (65) to second order $(\delta\tilde{h}_3)^2$. From (49) we get

$$\frac{1}{2}\,\Im[D_1(h) + D_2(h)] = -i\pi\rho(h), \quad h \in I_0, \tag{82}$$

which implies that

$$\rho(h) = \rho^{kz}(h; h_1, h_2, h_3; \alpha, \beta) + O((\delta\tilde{h}_3)^2), \tag{83}$$

where the explicit form of the $O((\delta\tilde{h}_3)^2)$-corrections can be read off from (80). An expansion of the constraints $\mathcal{W}_j$ to second order yields

$$\begin{aligned} \mathcal{W}_j &= W^{kz}(h_1, h_2, h_3, \alpha, \beta) \\ &\quad + \frac{(\delta\tilde{h}_3)^2}{4}\left(\frac{\partial F_3(h_3)}{\partial h_3} + \frac{F_3(h_3)}{2(h_3-h_1)} + \frac{F_3(h_3)}{2(h_3-h_2)})\right) + \cdots = 0. \end{aligned} \tag{84}$$

In a similar way we obtain explicit expressions for $\mathcal{O}_j$ to order $(\delta\tilde{h}_3)^2$.

These expressions can now be used to refine the relations (72) which govern the behaviour along an arbitrary, non-tangential approach to the critical line. One obtains

$$\begin{aligned} \delta\alpha \sim \delta\beta &\sim \delta h_2, \\ (\delta h_3)^2 + (\delta\tilde{h}_3)^2 &\sim a_1\delta\alpha + a_2\delta\beta \\ \delta h_3 \cdot \delta\tilde{h}_3 &\sim \delta\tilde{\alpha}, \end{aligned} \tag{85}$$

for some real a_i. These relations can be diagonalized to give

$$(\delta h_3 \pm \delta\tilde{h}_3)^2 \sim (\delta h_3^{(i)})^2 \sim a_1\delta\alpha + a_2\delta\beta \pm a_3\delta\tilde{\alpha}, \tag{86}$$

characterizing the critical behaviour related to the approach to the two critical surfaces discussed earlier.

By contrast, in phase B we have only a single smooth critical hypersurface characterized by $\mathcal{F}_2^{(j)}=0$. Again, moving tangentially to this surface will induce linear variations which are all of the same order, whereas generic, non-tangential approaches to the surface will be characterized by

$$\delta\alpha \sim \delta\beta \sim \delta\tilde{\alpha} \sim \delta h_3 \sim \delta\tilde{h}_3 \sim (\delta h_2)^2. \tag{87}$$

5.2 The scaling relations and renormalization of 3D gravity

We now want to relate the scaling limit of the matrix model discussed above to the continuum limit of the discretized 3D quantum gravity. (Euclidean) quantum field theories can be defined as the scaling limit of suitable discretized statistical theories. The continuum coupling constants are then defined by a specific approach to a critical point of the statistical theory. Different approaches to the critical point might lead to different coupling constants or even different continuum theories. In our case we want to show that it is possible to approach a critical point in such away that the canonical scaling expected from a theory of 3D gravity is reproduced.

We saw above that the asymmetric model is defined on a three-dimensional hypersurface in the parameter space of $(h_1, h_2, h_3^{(i)})$. Let us consider a curve approaching a specific symmetric point $(h_1^c, h_2^c, h_3^{(i)}=h_3^c)$ on the critical surface of the model, with curve parameter a (where $a=0$ corresponds to the critical point). In terms of independent parameters, this path can be described by $(h_2(a), h_3^{(1)}(a), h_3^{(2)}(a))$ or, equivalently, $(\alpha_1(a), \alpha_2(a), \beta(a))$. Recalling our discussion in Sec.2 above, we are interested in obtaining a behaviour of the form

$$\begin{aligned} \log\alpha_i &= \log\alpha_c + \frac{a\, c_1}{G} - a^2 Z_i - a^3 c_2 \Lambda \\ \log\beta &= \log\beta_c - \frac{a\, c_3}{G} - a^3 c_4 \Lambda \end{aligned} \tag{88}$$

as $a \to 0$. In (88), $\log \alpha_c$ and $\log \beta_c$ represent combined additive renormalizations of all the coupling constants, and G, Λ and Z_i are the renormalized gravitational and bulk and surface cosmological coupling constants.

For the purposes of conventional quantum gravity we are primarily interested in phase A of the model. Moreover, because of the symmetry between in- and out-states which is related to our three-dimensional geometric interpretation of the essentially two-dimensional matrix model, we are only interested in continuum limits where both of the discrete spatial boundary volumes go to infinity, that is, where both z_t and z_{t+a} are critical. Interesting critical points therefore must lie on the critical line where the two critical surfaces intersect.

As we have argued at length in [9], a standard critical behaviour of 3D gravity governed by the canonical scaling of the three-volume can only be achieved by a tangential approach to the critical line (in that case, in the β-α plane). This approach ensures that the terms proportional to a/G in the expansions of α and β by themselves do not determine the leading critical behaviour, although they contribute in the combination ΛG^3 at higher order. The same construction carries over to the asymmetric case, which differs from the symmetric situation by the appearance of the second-order terms proportional to Z_i in (88). As happened there, the tangent vector to the critical curve (in the hyperplane $\tilde{\alpha} = 0$) at a given point $\alpha^c(\beta)$ fixes the ratio c_1/c_3 of the constants appearing in (88). Moreover, the leading critical behaviour will now be governed by the boundary cosmological terms. This means in practice that in order to determine the singular behaviour of the free energy $F(\alpha_i, \beta)$ under such an approach, we can simply ignore the order-a terms in (88) and use the scaling relations (85) for the remainder.

We still have to discuss the relative scaling behaviour of $\delta\tilde{\alpha}$ and $\delta\alpha$. As we will show in the next section, if we want the transfer matrix to reduce to the unit operator in the limit $a \to 0$, the singular behaviour should to leading order depend only on $\delta\alpha$ and not on $\delta\tilde{\alpha}$. One way of realizing this possibility would be by showing that at this order the coefficient in front of the $\delta\tilde{\alpha}$-term vanishes, implying the relations

$$\begin{aligned} \delta\alpha &\propto a^2, \\ \delta\tilde{\alpha} &\propto a^3 \end{aligned} \tag{89}$$

and consequently

$$\begin{aligned} \delta h_3 &\propto a, \\ \delta\tilde{h}_3 &\propto a^2. \end{aligned} \tag{90}$$

The corresponding scaling curve deviates from the symmetric plane $\alpha_1 = \alpha_2$ only by $O(a^3)$-terms. Note that this means that the asymmetry $\delta\tilde{\alpha}$ contributes at the same order as the cosmological term Λ, which presumably is a desirable property in view of the construction of the Hamiltonian.

6 The transfer matrix

The transfer matrix contains the information necessary to derive the Hamiltonian of 3D gravity. From the free energy of the asymmetric ABAB model we can extract some information about the transfer matrix as is clear from formula (4). Let us be more precise about this (see [1] for a detailed discussion). The free energy of the asymmetric ABAB matrix model involves according to (4) a summation over the individual geometric states $|g\rangle$ which label in- and out-states. However, one can use the free energy to extract information about the areas N_{in} and N_{out} (the number of squares in the in- and out-quadrangulations) of the in- and out-states $|g_{in}\rangle$ and $|g_{out}\rangle$. We expect this quantity to capture the essential part of physical information about the time evolution of a two-dimensional universe (cf. e.g. [14]). Let us consider the state

$$|N\rangle = \frac{1}{\sqrt{\mathcal{N}(N)}} \sum_{g_t} \delta_{N,N(g_t)} |g_t > \tag{91}$$

where $\mathcal{N}(N)$ is the number of quadrangulations of given area N. The norm of such a state is

$$\langle N'|N\rangle = \delta_{N,N'} \tag{92}$$

since states $|g_1\rangle$ and $|g_2\rangle$ with different quadrangulations are orthogonal. The number of quadrangulations constructed from N squares grows exponentially as

$$\mathcal{N}(N) = N^{-7/2}\, \mathrm{e}^{\mu_0 N}(1 + O(1/N^2)), \tag{93}$$

where μ_0 is known. The sum (4) can now be written as

$$F(\alpha_1, \alpha_2, \beta) = \sum_{N_t, N_{t+a}} \mathrm{e}^{-z_t N_t - z_{t+a} N_{t+a}} \langle N_{t+a}|\hat{T}|N_t\rangle \sqrt{\mathcal{N}(N_t)\mathcal{N}(N_{t+a})}. \tag{94}$$

The exponential part of $\sqrt{\mathcal{N}(N_t)\mathcal{N}(N_{t+a})}$ can be absorbed in additive renormalizations of the boundary cosmological constants z_t and z_{t+a} (i.e. additive renormalizations of $\log \alpha_i$, recall (4)). It follows that in the scaling limit, i.e. for large N where we can use (93), the Laplace transforms of the matrix elements $\langle N_1|\hat{T}|N_2\rangle$ are equal to the "7/2" fractional derivative[7] of the free energy $F(\alpha_1, \alpha_2, \beta)$,

$$\sum_{N_t, N_{t+a}} \mathrm{e}^{-z_t N_t - z_{t+a} N_{t+a}} \langle N_{t+a}|\hat{T}|N_t\rangle = \left(\frac{\partial}{\partial z_t} \frac{\partial}{\partial z_{t+a}}\right)^{7/4} F(\alpha_1, \alpha_2, \beta). \tag{95}$$

where $-\frac{\partial}{\partial z} = \frac{\partial}{\partial \log \alpha}$ as is clear from (5).

The scaling limit, and thus the continuum physics, is determined by the singular part of the free energy. The leading behaviour of this singular part when we approach

[7] There are standard ways to define the concept of a fractional derivative, see for instance [13].

a critical point as described in the previous section, is given by $F(\alpha_1, \alpha_2, \beta) \propto (\delta\alpha)^{5/2}$. It is now straightforward to apply (95) and one finds

$$\left(\frac{\partial}{\partial \log \alpha_1}\right)^{7/4} \left(\frac{\partial}{\partial \log \alpha_2}\right)^{7/4} F(\alpha_1, \alpha_2, \beta) \approx \frac{1}{\delta\alpha}. \tag{96}$$

This is exactly the leading-order behaviour we expect for the transfer matrix when $a \to 0$ from (1):

$$\langle N_1|\mathrm{e}^{-a\hat{H}}|N_2\rangle \to \langle N_1|\hat{I}|N_2\rangle = \delta_{N_1,N_2}, \tag{97}$$

and the Laplace transform of δ_{N_1,N_2} is (for large N's)

$$\frac{1}{\delta\alpha_1 + \delta\alpha_2} = \frac{1}{a^2(Z_1 + Z_2)}(1 + O(a)). \tag{98}$$

It would be very interesting to expand to next order in a and thus obtain information about $\hat{H}$. These terms come from the $O(a^3)$-terms discussed in the previous section. This will only give us information about the matrix elements related to the states of the form (91), but as discussed in detail in [1] we expect that this is the only information relevant in the continuum limit of 3D gravity if the topology of the spatial slices is that of a sphere.

Acknowledgement

All authors acknowledge support by the EU network on "Discrete Random Geometry", grant HPRN-CT-1999-00161. In addition, J.A. and J.J. were supported by "MaPhySto", the Center of Mathematical Physics and Stochastics, financed by the National Danish Research Foundation. J.J. acknowledges support by Polish Committee for Scientific Research (KBN) grant 2 P03B 096 22 (2002-2004).

References

[1] J. Ambjørn, J. Jurkiewicz, R. Loll and G. Vernizzi: *Lorentzian 3D gravity with wormholes via matrix models*, JHEP 0109 (2001) 022 [hep-th/0106082].

[2] J. Ambjørn, J. Jurkiewicz and R. Loll: *A nonperturbative Lorentzian path integral for gravity*, Phys. Rev. Lett. 85 (2000) 924-927 [hep-th/0002050].

[3] J. Ambjørn, J. Jurkiewicz and R. Loll: *Dynamically triangulating Lorentzian quantum gravity*, Nucl. Phys. B 610 (2001) 347-382 [hep-th/0105267].

[4] J. Ambjørn and R. Loll: *Non-perturbative Lorentzian quantum gravity, causality and topology change*, Nucl. Phys. B 536 (1998) 407-434 [hep-th/9805108].

[5] J. Ambjørn, J. Jurkiewicz and R. Loll: *Nonperturbative 3-d lorentzian quantum gravity*, Phys. Rev. D 64 (2001) 044011 [hep-th/0011276].

[6] J. Ambjørn, J. Jurkiewicz and R. Loll: *Computer simulations of 3d Lorentzian quantum gravity*, Nucl. Phys. Proc. Suppl. 94 (2001) 689-692 [hep-lat/0011055]; *3d Lorentzian, dynamically triangulated quantum gravity*, Nucl. Phys. Proc. Suppl. 106 (2002) 980-982 [hep-lat/0201013].

[7] V.A. Kazakov and P. Zinn-Justin: *Two matrix model with ABAB interaction.*, Nucl. Phys. B 546 (1999) 647-668 [hep-th/9808043].

[8] V.A. Kazakov, M. Staudacher and T. Wynter: *Exact solution of discrete two-dimensional R^2 gravity*, Nucl. Phys. B 471, (1996) 309-333 [hep-th/9601069]; *Almost flat planar diagrams*, Commun. Math. Phys. 179 (1996) 235-256 [hep-th/9506174]; *Character expansion methods for matrix models of dually weighted graphs*, Commun. Math. Phys. 177 (1996) 451-468 [hep-th/9502132]; P. Zinn-Justin: *Random hermitian matrices in an external field*, Nucl. Phys. B 497 (1997) 725-732 [cond-mat/9703033].

[9] J. Ambjørn, J. Jurkiewicz and R. Loll: *Renormalization of 3d quantum gravity from matrix models*, [hep-th/0307263].

[10] P. Zinn-Justin: *The asymmetric ABAB matrix model* [hep-th/0308132].

[11] J. Ambjørn, J. Correia, C. Kristjansen and R. Loll: *The relation between Euclidean and Lorentzian 2D quantum gravity*, Phys. Lett. B 475 (2000) 24-32 [hep-th/9912267].

[12] N.I. Muskhelishvili: *Singular Integral Equations*, Dover Publn. 1992.

[13] S. G. Samko, A. A. Kilbas and O. I. Marichev: *Fractional integrals and derivatives: theory and applications*, Gordon & Breach, 1987.

[14] S. Carlip: *Quantum gravity in 2+1 dimensions*, Cambridge Univ. Press, Cambridge, UK (1998).